R E P O R T R E S U M E S

ED 003 078 24

DEVELOPMENT AND VALIDATION OF DESCRIPTIVE SCALE FOR MEASUREMENT OF ART PRODUCTS.

BY- ROUSE, MARY J.
INDIANA UNIV., BLOOMINGTON

REPORT NUMBER CRP-S-077 PUB DATE 65
CONTRACT OEC-4-10-227
EDRS PRICE MF-$0.09 HC-$1.88 47P.

DESCRIPTORS- *CREATIVITY RESEARCH, *RATING SCALES, *ART PRODUCTS, *MEASUREMENT TECHNIQUES, CREATIVE ART, *MEASUREMENT INSTRUMENTS, BLOOMINGTON, INDIANA

IDENTIFICATION AND DEFINITION OF TRAITS AND CHARACTERISTICS USUALLY PRESENT IN MOST KINDS OF ART PRODUCTS WERE ATTEMPTED DURING THE REPORTED RESEARCH. THE STUDY SOUGHT TO COMBINE THE IDENTIFIED AND DEFINED CRITERIA INTO A USABLE SCALE AND TO TRAIN JUDGES (GRADUATE ART EDUCATION MAJORS) IN THE USE OF SUCH A SCALE. PAINTINGS, DRAWINGS, AND CONSTRUCTIONS WERE COLLECTED FROM ELEMENTARY SCHOOL CHILDREN, JUNIOR AND SENIOR HIGH SCHOOL STUDENTS, AND UNIVERSITY-LEVEL NONART AND ART EDUCATION MAJORS. SAMPLES WERE DRAWN FROM THIS POOL IN RANDOM FASHION FOR USE IN THE STUDY. THESE ART PRODUCTS WERE JUDGED ON THE BASIS OF A PROVISIONAL SCALE SELECTED FROM RESULTS OF A SURVEY OF EXPERTS IN THE FIELDS OF ART, ART EDUCATION, AND ART CRITICISM. A REFINED SCALE WAS DEVELOPED AS A RESULT OF A STATISTICAL ANALYSIS BY WHICH IT SHOULD BE POSSIBLE TO MEASURE THE QUALITY OF ART PRODUCTS. (WB)

ED003078

U. S. DEPARTMENT OF HEALTH, EDUCATION AND WELFARE
Office of Education
This document has been reproduced exactly as received from the person or organization originating it. Points of view or opinions stated do not necessarily represent official Office of Education position or policy.

DEVELOPMENT AND VALIDATION OF DESCRIPTIVE SCALE FOR MEASUREMENT OF ART PRODUCTS

Cooperative Research Project No. S-077
OE-4-10-227

1965

Mary J. Rouse

Indiana University
Bloomington, Indiana

The research reported herein was supported by the Cooperative Research Program of the Office of Education Department of Health, Education, and Welfare

U. S. DEPARTMENT OF HEALTH, EDUCATION AND WELFARE
Office of Education
This document has been reproduced exactly as received from the person or organization originating it. Points of view or opinions stated do not necessarily represent official Office of Education position or policy.

Accepted
4-7-65
S-077

Table of Contents

	Page
1. The Problem	4
A. Significance of the Problem	4
2. Objectives	5
3. Related Research	6
4. Procedures	7
5. Analysis of the Findings and Results	8
A. Collection of the Samples of Art Products	8
B. The Development of the Scale	9
C. The Refinement of the Scale	10
Reliability Analysis	12
Validation Analysis	13
6. Conclusions	14
A. Judging and Reliability Analysis	14
B. The Validation Analysis	16
7. Implications	21
8. Bibliography	23

LIST OF TABLES

Table		Page
1.	38 Items Submitted to Experts for Rating in Terms of Possible Efficiency	25
2.	Letter Sent to Experts Concerning the Rating of the 38 Items	26
3.	Instruction Sheet for Rating the 38 Items	27
4.	Results of Rating by Experts in Terms of Median and Q-Scores (A Measure of Ambiguity of Ratings	28
5.	Interjudge Correlations, Communalities, Means and Standard Deviations	29
6.	Intercorrelations of Criteria Average Scores on 20 Variables	30
7.	Results of Factor Analysis Showing Six Factors and 20 Criteria	31-32
8.	Results of T-tests of Mean Scores Achieved by "Best" and "Poorest" Products	33
9.	Final Version of the Measure	34-43

1. The Problem

This study attempted to develop and to establish the reliability and validity of a descriptive measure for use in future research studies in the field of art education. It was hoped that this measure could identify the various descriptive characteristics which usually appear in different kinds of art products. It was hypothesized that such an instrument would be extremely useful in assessing changes of artistic behavior over periods of time and under different treatments.

The problem involved several aspects:

a. the identification of descriptive characteristics which appear in several kinds of art products, the delimitation and definition of these characteristics, and the selection of the most adequate of these for inclusion in the measure.

b. the collection of a suitable sample of art products to use in the development of the measure.

c. the training of judges in proper rating procedures.

d. the determination of the reliability and validity of the measure.

A. Significance of the Problem

Although artists and aestheticians have debated about works of art for centuries, there has been very little real agreement among them concerning the descriptive characteristics to be used in such discussions. Each writer appeared to believe that his, and only his, individual list of terminology was suitable for the task. Further, each of these terms usually was defined as the individual chose with little reference to prior definitions. Thus minimal communication resulted. Although writers in these fields have issued a steady stream of books and articles, each was largely unrelated to others, and usually turned out to be an argument in favor of an individual theory. Only scattered attempts have been made to draw these viewpoints together into any sort of integrated structure, and consequently, few real advances have been made. The field of art criticism remains, therefore, still in an ambiguous state which can best be described as a highly esoteric debating society.

Voices, of course, have called for reform and structure at regular intervals but have usually been ignored. Thomas Munro, for example, wrote, "There is no obstacle but the inertia of tradition to prevent aesthetics from undertaking an extensive program of direct comparative observation of particular examples from various arts, with the aim of discovering common and divergent qualities of form."[6]

In spite of Munro's statement, however, little direct comparative observation has been forthcoming in that field.

While this nebulous state may be sufficient for aestheticians and possibly studio artists, it clearly is not adequate for empirical research in art or for the teaching of art in the public schools. In both of these areas, clarification of terminology and identification of those phenomena which can be described and quantified are matters of first importance.

Psychologists and researchers from the field of art education have developed the skills that will allow them to design measures which can deal with these phenomena and some attempts have already been made to analyze certain aspects of art products.

Usually, however, these efforts have been concerned only with special characteristics in which their designers had some personal interest, rather than more generalized art characteristics. There have been, for example, several scales concerned with stylistic orientation. Others have dealt with a construct termed "spontaneity" and several have been concerned with the investigation of various aspects of "creativity."

Most of these scales, however, were developed in a deductive manner; that is: they were based on the investigators' prior theoretical position and reflected these in various limitations. The inductive manner was rarely used, and because of this, it is possible that many criteria may have been overlooked and hence not investigated.

For these reasons, it seemed important that a study should be carried out solely for the purpose of revealing as many efficient criteria or characteristics as could be found in several kinds of art products without regard as to whether these might confirm or deny some theoretical construct. It also seemed important to attempt to define these characteristics operationally and to include them in a scale that could be utilized for future studies concerned with changes in expressive graphic behavior.

2. Objectives

Designers of proposed measures of art products must always face the question of whether normative statements concerning the "goodness" or "badness" of products can be made. Aestheticians have debated this question for centuries and have not yet arrived at a decision as to whether this can be accomplished, and if so, how it should be accomplished. This study attempted to avoid this evaluative issue and sought instead to identify and utilize those characteristics which might be termed "descriptive."

In order to accomplish this, the following objectives were stated:

A. The study would seek to identify and define operationally those traits or characteristics usually present in most kinds of art products. This objective eliminated for consideration those possible criteria dealing with aspects of color, which may or may not be present in art products.

B. The study would seek to avoid criteria related to "goodness" or "badness" of the art product.

C. The study would seek to combine the identified and defined characteristics into a usuable scale and to train judges in the use of it.

D. The study would seek to assess the reliability and validity of the final measure by the process of having three judges use it to independently rate a heterogeneous range of art products of three kinds: drawings, paintings, and three-dimensional constructions.

3. Related Research

M. H. Bird attempted to design a measuring instrument for drawings in order to test the relationship of intelligence to drawing ability.[3]

Kline and Carey sought also to establish a measuring instrument for free-hand drawings, using criteria based on a theory of the characteristics thought necessary for good design.[13]

G. M. Kinget designed a test using an analysis of the formal characteristics of drawings as a basis for diagnostic psychological statements.[12] Karen Machover also designed another drawing test for the same use.[14] Florence Goodenough brought forth a test for children's drawings which she felt could determine the children's I.Q. levels.[7] More recently, D. B. Harris restandardized the Goodenough drawing test, but limited it rather severely in terms of possible usage as a test of intelligence.[11]

Elkisch utilized some formal criteria in an analysis of paintings in an attempt to provide psychological diagnostic material.[6]

Wolff developed a "Table of Graphic Elements" in order to better evaluate child personality.[20]

In the art education field, several studies of interest have been conducted recently. June McFee sought to describe different kinds of art products in terms of their degree of creativeness through the assessment of observable qualities thought to be related to creativity.[15]

E. W. Eisner attempted also to measure two disimilar kinds of art products in terms of creative criteria. His study included both drawings and constructions.[5]

R. Burkhart designed a scale which was thought to assess the degree of "spontaneity or deliberateness" present in paintings. By his own definition, this dichotomy was equated with the presence or lack of creativity.[4]

Gogel produced an eight-point scale which measured this same dichotomy in drawings.[4] Interested in the same theory, Nitschke designed a test which counted the blackness of lines within drawings. He believed that the blacker the lines, the more spontaneous and hence, creative, the drawing.[17]

Beittel and R. Burkhart published a study in 1963, entitled "Strategies of Spontaneous, Divergent, and Academic Art Students," which utilized a scale to measure drawings. This study enlarged and expanded Burkhart's previous construct of "spontaneity" and "deliberateness" as representative of degrees of creativity.[1]

In a report published some months after the investigator's own study was begun, Beittel and Burkhart described another of their investigations which sought to isolate those formal elements or characteristics in art products which might be part of the construct they had previously labeled "spontaneity." These researchers inspected a sample of paintings in an attempt to identify those "criterion elements" of spontaneity which clearly were present. Of particular interest to the present study, they tried to discover characteristics "so elementalistic that only a brief title would be needed to convey its definition to the judge and would leave no room for the judge to define it in his own terms."[2] They isolated six elements: "detail," "media overlap," "border of shapes," "texture," "movement within shapes," and "moving objects." They next attempted to isolate so-called "formal elements" and listed them as: "overlap-transparency," "progressive abstraction," "connectedness," "three-dimensionality" and "dominance-subordination." They also developed a third group which they labeled "divergent." These criteria were: "alteration of viewpoints," "rearrangements," and "alteration of treatments."

A factor analysis was conducted of their results and it was found that some of the elements were intercorrelated as had been hypothesized and some were not. It will be shown later that a number of those criteria are similar to those identified by our investigation while others are not similar.

4. Procedure

The procedure that was followed in this study was to (a) assemble a collection of possibly relevant items for inclusion in the scale from all available sources, including the studies already cited and the writings of aestheticians, art historians, designers, and art

educators; (b) to delimit these items to those most often specified by these writers; (c) to define the remaining items provisionally; (d) to submit these possible items to forty-eight experts in the fields of art, art education, and art criticism, asking each to rank them according to which should prove most efficient for the purposes of this study; (e) to analyze the returns in order to select those most chosen by these experts; (f) to redefine these selected items in operational terms; (g) to collect art products in three media (paintings, drawings, and constructions) from children in elementary, junior and senior high school classes and from some students at the university level; (h) to train judges in the use of the preliminary measure; (i) to conduct a preliminary judging of a sample of the products in an attempt to eliminate those which could not be used in reliable fashion; (j) to construct a final scale after eliminating the unreliable items and in accordance with certain findings relating to necessary judging requirements; (k) to use the judges and the redesigned scale in an assessment of 50 art products; (l) to analyze the results in terms of reliability; and (m) to subject the results to a factor analysis and other procedures in an attempt to establish validity.

5. Analysis of the findings and results

A. Collection of the sample of art products

Over a period of two months the investigator and assistant collected a sample of three kinds of art products: drawings, paintings, and three-dimensional constructions, from various levels of the Indiana University laboratory school. These levels were the second grade, the fourth grade, the seventh grade, and a senior high art class which included students from grades nine through twelve. Some products were also collected from university students, art education majors and non-majors alike.

Materials and stimulation for all these classes were provided by the investigator and her assistant. Materials were similar for all age levels. For the drawings, soft pencils, erasers, and sheets of standard-size white drawing paper were supplied. The painting materials included brushes of varied sizes, cake tempera of assorted colors, and sheets of white paper. The constructions were made from balsa wood strips of several sizes cut to measure by the individuals and assembled with model cement and pins. Colored string and construction paper were also available for use.

After the collection of products was completed, they were assembled into one pool. Equal numbers of drawings, paintings, and constructions were randomly selected from the pool and given numbers which were henceforth to be the sole means of identification. One-half of the pool was used for training the judges and refining the scale. The other half was reserved and used for the final judging process on which the analyses of reliability and validity were based.

It is felt that this procedure offered an extremely heterogeneous collection of products for the judging process and that this contributed an added degree of difficulty which most other studies which dealt with art products did not possess.

B. The development of the scale.

The first step in the isolation of possible descriptive items was the consultation of the writings of authorities in the fields of aesthetics, art history, design, and art education, as well as prior empirical studies dealing with any aspect of graphic art. Each term mentioned in these sources was recorded and a check was kept on the number of times each term was used by different experts. Some 118 books and studies were consulted over a period of two months and a total of 313 different terms were discovered. Some of these were synonyms or antonyms, but the length of the list was sufficient to show the rampant confusion concerning terminology prevailing in these fields. Clearly, very little agreement, and therefore, communication, had been established over a period of generations.

Since 313 terms obviously could not be utilized in the proposed scale, some method of limitation had to be found. The investigator, therefore, decided to consider only those terms mentioned by five writers or more. This procedure resulted in a list of 51 terms which represented 16% of the larger list. Since the objective had been to use only those characteristics which might be found in several kinds of art products, some of these could be eliminated without difficulty. Others were obvious synonyms or antonyms and were also discarded. By these means, the list was cut to 38 items. These are listed in the Appendix.

The next step was to attempt to define these terms operationally. The assumption was that the 118 books and studies already consulted would provide adequate definitions, but this did not prove to be the case. Not only were these definitions usually not employed in a standardized manner by the various writers, but the definitions were so nebulous and ambiguous that few were of any use in this study. Further help was sought from dictionaries of all types, as well as any other source which might be thought helpful in an enterprise of this nature, but they were all of only minor assistance.

The decision finally was to use the best definitions available at the moment for the rating procedure with the qualification that they were not final. Criticism and suggestions for revisions were requested from the experts consulted in the rating procedure conducted next.

The 38 items and their temporary definitions were dispatched to 48 experts from the fields of art education, art history, design, and aesthetics with the request that they rank them using a Q-sort method, according to their efficiency for the purposes specified. The directions for the Q-sort are included in the Appendix.

Replies from 39 experts were received in time to be used in the analysis. Four of these had to be rejected because of errors and five were outright refusals to participate in the project. Three of the refusals came from individuals who felt they did not have time to cooperate but the other two are worth commenting on because their reasons seem typical of the attitudes of many individuals within these fields. One respondent wrote that he did not wish to participate because studies of this nature were attempting to make artistic creation too scientific. The other felt that this kind of study simply could not be done.

Other experts, not only participated in the rating operation but also furnished the requested criticism of the definitions and this information was most helpful in the redefining process which took place.

The returns were analyzed using a system recommended by Goode and Hatt.[8] Cumulative graphs were plotted for each item and medians and quartile values established. The median was used as one score and the "Q" figure, which was the difference between the first and third quartiles, was also used since this represented an indication of the ambiguity of the item. Table 4 in the appendix shows the results of this analysis, which resulted in the selection of 28 items for preliminary judging activity.

C. The refinement of the scale.

With the selection of the 28 items, the first version of the scale could be constructed. Each item followed the form:

> "Is there any evidence of unity? (Unity would be shown by the object's possessing a sense of oneness, i.e.: all means and elements appear to be adapted to a single purpose or end.)
>
> None Some More Much

The judges, two male and one female art education graduate students, were brought together at intervals to practice using the preliminary scale in judging small numbers of all three kinds of art products. The investigator worked closely with this group as they practiced together,but each judge rated the products independently without the presence of the investigator.

Reliability checks were made at intervals as the judging preceeded. It soon became evident that "none," "some," "more," and "much" were not efficient quantifying terms. It was difficult for the judges to agree as to the meaning of these words.

These words were changed to numbers but this method also proved to be too ambiguous for the judges to utilize efficiently. Iterrater reliability coefficients were rarely above .30 at this stage.

Some items were shown to be strongly non-efficient or were felt by the judges to overlap with others and so were eliminated from the next series. These were: "Accentuation," "Proportion," and "Opposition." Even after they were eliminated, however, difficulty was still being encountered by the judges.

It became clear that some of these 28 items were dichotomies; that is: one item, "Simplicity" was actually the opposite of another, "Complexity." Similarly, "Spontaneity of Handling" was the opposite of "Deliberateness in Handling." Yet the judges were experiencing difficulty in scoring them in correct directions while separate. It seemed necessary to combine these sets of opposites into single dichotomous items.

All opposites were cast into dichotomies, therefore, and those items which were not obviously related to other items were restated into dichotomous terms. A vertical form was also adopted instead of the previous horizontal one, thus allowing for a specific definition of each of the steps in each item.

This new form forced a change in terminology involving some of the major characteristics which the scale purported to measure. Thus "Amount of Detail" was chosen as a better expression of the totality represented on one end by "Simplicity" and on the other by "Complexity." Also "Handling" seemed inclusive enough for both "Spontaneous Handling" and "Deliberate Handling." This item evolved, finally into "Carefulness" at one extreme and "Freeness" at the other. Those items which had dealt with "Similarity" before assumed a new identity as "Non-variation."

Since this combination process reduced the total number, it was thought that two additional items might well be introduced into the scale and tested empirically. These two, one termed "Mode: Realism/Abstraction" and the other labeled "Non-transparency/Transparency" had not been rated highly in terms of possible efficiency by the experts but it was decided that they should be tried with the proviso that elimination would follow should interrater reliability figures be low.

The judges also indicated that they felt an intuitive need for a center-step for each of the items (only four steps had been used in the previous models of the scale.) Therefore two new forms of the measure were designed, one having three steps and the other having five. These steps were specifically defined in each case.

The judges used first the three-step form and then the five-step model in rating different samples of the art products. Interrater reliability checks showed that the five-step form was superior to the one which had only three steps and so the decision was made to employ it in the next rating series.

The judges, using the redefined five-step scale, then rated the final 50 art products, and the results of this judging were used for the final reliability and validity analyses. A copy of the scale is included in the Appendix.

D. Reliability Analysis

An analysis was made of interrater agreement on each of the 20 items since it had become clear that these items were tapping different dimensions of artistic phenomena and therefore an analysis utilizing only one global score for the entire measure seemed illogical. The procedure used for these analyses was a two-way classification analysis of variance technique recommended by Guilford.[9] The interrater reliability coefficients obtained for each criteria is reported below.

Criteria	Correlation Coefficient
1. Shape: Non-variation/Variation	.70
2. Mass: Non-variation/Variation	.68
3. Line: Non-variation/Variation	.72
4. Amount of Texture	.91
5. Amount of Detail	.86
6. Flatness/Depth	.84
7. Non-tension/Tension	.61
8. Non-symbolism/Symbolism	.71
9. Amount of Balance	.63
10. Amount of Unity	.84
11. Non-dominance/Dominance	.66
12. Non-craftsmanship/Craftsmanship	.89
13. Line: Thinness/Thickness	.85
14. Non-gradation/Gradation	.85
15. Handling: Carefulness/Freeness	.85
16. Symmetry/Asymmetry	.63
17. Non-rhythm/Rhythm	.69
18. Non-transparency/Transparency	.88
19. Non-originality/Originality	.82
20. Mode: Realism/Abstraction	.83

It had been proposed originally that a split-half reliability figure should be calculated for the entire scale as another check on reliability. The results of the factor analysis of the scores, soon to be discussed, showed that this type of analysis was not appropriate because several different dimensions were clearly evident.

Another kind of reliability assessment, however, was made. After a period of time had elapsed between the first rating of the 50 products, a sample of these were chosen for re-rating. Each of the three judges rerated a number of the same products and the old and new scores were compared.

Judge A's reratings correlated .97 with the first ones, Judge B's correlated .89, and Judge C's reratings correlated .86 with the original series.

E. Validation Analysis

After the reliability figures were computed and the results proved adequate for retention of all 20 criteria, the investigator decided to subject the criteria to a more involved validation process than had been originally planned. Therefore a factor analysis of the data was conducted using a principal components program with varimax rotation.

This process clearly revealed that six factors were present within the 20 criteria. The results of the complete analysis, showing the loading of the factors on each of the 20 criteria is presented in Table 7 in the appendix.

These results can be summarized briefly.

"Shapes: Non-variation/Variation," "Mass: Non-variation/Variation," "Line: Non-variation/Variation," "Amount of Texture," and "Amount of Detail" were all loaded highly in a negative direction with Factor A which seems to relate to some aspects of non-differentiation. "Line: Thinness/Thickness," "Symmetry/Asymmetry," "Non-originality/Originality," "Amount of Gradation," and "Non-transparency/Transparency" also had some fairly heavy negative loading on this factor.

"Flatness/Depth," "Non-tension/Tension," and "Non-symbolism/Symbolism" had high positive loading on Factor B which apparently relates to space. "Non-originality/Originality," "Non-dominance/Dominance," and "Non-transparency/Transparency" had minor loadings on this factor also.

"Amount of Balance," "Amount of Unity," "Non-craftsmanship/Craftsmanship," and "Non-dominance/Dominance," were heavily loaded on Factor C which must represent a kind of structural-intellectual domain. "Handling: Carefulness/Freeness," "Non-symbolism/Symbolism," and "Symmetry/Asymmetry" had minor negative loadings on this factor.

"Handling: Carefulness/Freeness," "Line: Thinness/Thickness," and "Amount of Gradation" had high loadings on Factor D, which may reflect a kinesthetic-spontaneity construct. "Shapes: Non-variation/Variation," "Amount of Texture," and "Non-transparency/Transparency" were also lightly loaded with this factor in a positive direction while "Amount of Unity," "Non-craftsmanship/Craftsmanship," and "Non-dominance/Dominance" revealed light negative loadings.

"Symmetry/Asymmetry," "Non-rhythm/Rhythm," and "Non-transparency/Transparency" were shown to have heavy negative loadings on Factor E, while "Flatness/Depth," "Non-tension/Tension," "Non-craftsmanship/Craftsmanship," "Amount of Detail,"and "Amount of Gradation" showed smaller negative loadings. "Line: Thinness/Thickness," and "Amount of Balance" were revealed to have small positive loadings. It is thought that this factor may express the opposite of internally dynamic action, static action.

"Non-originality/Originality," "Mode: Realism/Abstractness," and "Non-tension/Tension," were shown to be heavily loaded, negatively, with Factor F, which might be termed conformity. "Non-craftsmanship/Craftsmanship," and "Non-rhythm/Rhythm" were more lightly loaded in a negative direction and "Amount of Gradation" in a positive direction on this factor.

As another check on validity, two faculty members in the art education department chose eight products which they considered globally "best" and eight products which they considered globally "poorest" from the entire sample of 50 products. Our judges' previous ratings of these 16 products were analyzed in terms of the six sub-test scores: and the "best" group's scores were compared with the "poorest" group's scores by t-tests. A significant difference ($p = .05$) was obtained between the scores of the two groups on one sub-test: Internal Static/Dynamic Action. The complete results are shown in Table 8.

6. Conclusions

A. Judging and Reliability Analysis

(1) Adequate reliability coefficients were achieved with all 20 criteria.

The fact that all 20 of the criteria proved empirically to have adequate interrater reliability levels seems enough justification for including all of them in the final scale. It is interesting that the two criteria which had been rejected by the experts in the previous validation-rating step, "Transparency" and "Mode: Realism vs. Abstraction," both showed very strong interrater agreement. "Transparency" achieved a reliability coefficient of .88 and "Mode" showed an .83 coefficient. This finding leads the investigator to believe that future studies should proceed directly toward empirical validation of possible criteria without recourse to "expert" opinion.

One of the weaker criteria, "Dominance" (.66) was also utilized in the last Beittel-Burkhart study.[2] Interrater correlation figures for this item in their study varied from .53 to .85.

The second type of reliability analysis made, a rerating process where ratings made by each judge of several products were compared with the same judge's reratings of these products after a time lapse of several days, showed adequate levels. Judge A's reratings correlated .97 with the first ratings, Judge B's reratings correlated .89 and Judge C's reratings correlated .86 with the first ones.

(2) Specifically defined criteria and quantifying steps were found to be necessary for successful rating processes.

Part of the success of the rating operations probably resulted from the investigator's attempt to structure the definitions as explicityly as possible. This included the clarification of the five steps quantifying each of the criteria. The judges themselves had indicated that this specificity might make their rating process more efficient, and had also requested a central quantification step. This reliance on intuitively-based procedures may well have important implications for explorations of a similar nature in the future, especially in the development of instruments in a field such as this where so little is currently known.

(3) The development of better working definitions for common terminology is a matter of primary importance.

The difficulties encountered by the investigator in attempting to pin down specific and workable definitions for our criteria needs reemphasizing. Clearly, the current lack of agreement between authorities on terminology should be a matter of utmost concern to everyone in these fields. It is the investigator's opinion that this task belongs logically to those groups who call themselves philosophers of art or aestheticians rather than to persons engaged in art education or psychological research. Yet, just as clearly, neither of these groups has expended much energy in this direction. Because of their avoidance of such important and basic responsibilities, these fields have shown little progress in past years and probably will continue to show as little in the future unless considerable reorientation takes place. In the meantime, others who are not content with the situation are faced with the necsssity of working out definitions as best they can.

This writer would be the first to agree that the definitions used for the criteria in this study are not the ultimate answers. It is conceivable that future studies may bring forth better and more specifically delimited ones, and indeed, this would be our fondest wish. The fact remains, however, that deficient as these definitions may be, they have served efficiently for the purpose at hand as our

reliability analysis showed. They should be considered, therefore, to be as or more valid than most others proposed before.

Further progress in investigations of this nature can only come on the basis of acceptance as to common meanings of important terminology. Some foundations for agreement may have been laid by this study and the investigator hopes that others might take these results and go forward.

B. The Validation Analysis

(1) One global test could not be considered adequate.

The factor analysis confirmed the suspicions voiced by the investigator and judges during the rating process that the 20 criteria were tapping several different dimensions of artistic phenomena. An earlier proposal had been made to combine all the criteria into a single test with one general score resulting, but the analysis showed that this would not be practical. Clearly, one global score would not reflect the several kinds of behavior changes which might otherwise be described.

(2) The 20 criteris, by means of a factor analysis, were found to belong to six different domains.

These 20 criteria actually appear to organize themselves into six different domains or constructs: (1) one concerned with a lack of differentiation, (2) a second involving spatial aspects, (3) another concerned with intellectual structuring, (4) one involving kinesthetic-spontaneity, (5) a fifth which probably describes internal static action, and (6) a final one which reflects conformity.

(3) It was necessary to redesign the criteria into six sub-tests, each comprising one factor, instead of the single global test proposed earlier.

Because of the results of the factor analysis, the decision was made to structure the 20 criteria into six sub-tests, each of which focussed on one of the factors and would provide a separate total score. This could be accomplished quite easily since the varimax rotation process tends to "clean up" factors, and makes it possible to determine those specific criteria which most obviously belong together.

The six sub-tests consist of these factors and criteria:

1. Test One: Non-differentiation/Differentiation

    a. Shapes: Non-variation/Variation
    b. Masses: Non-variation/Variation
    c. Lines: Non-variation/Variation
    d. Amount of Texture
    e. Amount of Detail

2. Test Two: Space

    a. Non-depth/Depth
    b. Non-tension/Tension
    c. Non-symbolism/Symbolism

3. Test Three: Structural/Intellectual

    a. Non-balance/Balance
    b. Non-unity/Unity
    c. Non-craftsmanship/Craftsmanship
    d. Non-dominance/Dominance

4. Test Four: Kinesthetic/Spontaneity

    a. Weight of Line
    b. Non-gradation/Gradation
    c. Handling: Carefulness/Freeness

5. Test Five: Internal Static/Dynamic Action

    a. Symmetry/Asymmetry
    b. Non-rhythm/Rhythm
    c. Non-transparency/Transparency

6. Test Six: Conformity/Non-conformity

    a. Non-originality/Originality
    b. Mode: Realism/Abstraction

The final scale, to be found in the Appendix, has been organized in this form and it is recommended that this version should be the one to be used in future studies. Separate scores can be computed for each sub-test and utilized for analyses, rather than calculating one global score. This procedure should make possible more precise descriptions of behavior changes.

(4) Three factors emerged because of negative relationships.

It is interesting to note that three of these factors were discovered in a negative fashion: that is, the

scores of the criteria making up these factors generally ran in an opposite direction from those of the other criteria. This negative direction posed some questions for our attempt to name these factors. It seemed logical, however, to give them names denotative of their negative characteristics, and so they were termed "Non-differentiation," "Static Motion," and "Conformity."

All three factors seem to be logically explainable. We recall that this sample of art products came mainly from youngsters of several grade levels in a public school. These individuals were not selected for special artistic talents, or for qualities of creativity. They certainly cannot be classified as mature persons.

"Differentiation" is considered to be an important characteristic of the mature individual by Witkin who has linked perception with personality characteristics.[19] Heinz Werner, in his investigations of symbol formation, also regards "differentiation" as highly important in the assessment of maturity of development.[18] Since our study deals with products of immature youngsters, it seems only logical to expect to find a fairly low level of "differentiation" appearing in our sample.

Conversely, "Conformity" might be expected to appear with frequency in products of these youngsters. Its opposite quality—"Uniqueness," by definition alone, would not be expected to show up as often.

The third factor which we termed "Static Action seems somewhat related to the other two in terms of expectancy alone, if for no other reason.

If, as we suspect, these three factors have some special relationship in common (since all three were revealed by negative correlations), we might hypothesize that this linkage may have important implications for future studies. Let us here consider each of these factor's opposite nature: instead of "Non-differentiation" substitute "Differentiation;" instead of "Static Action" read "Dynamic Action;" and instead of "Conformity" substitute "Uniqueness." With the realization that none of these opposite characteristics should be expected to appear with any frequency in the art products of average public school students (and did not do so in this study) it might be possible that we have isolated some of the factors composing that elusive entity we call "Creativity." This is, of course, only a conjecture but it bears further investigation.

(5) A fourth factor may be similar to one identified in other research studies.

Appropriate here is a discussion of one of the other factors which was found to exist in an appreciable quantity in our sample: the factor we termed "Kinesthetic-spontaneity." Admittedly, we have borrowed part of this term from prior research in art education reported by Beittel and Burkhart.[1, 2, 4] These investigators believed that "Spontaneity" represented an essential characteristic of creativity in the visual arts and conducted a series of studies concerned with it in one way or another. The factor we have chosen to term "Kinesthetic-Spontaneity" bears a distinct resemblance to that aggregate of characteristics that Beittel and Burkhart, and others working with them, described in the two earlier reports of their studies dealing with creative types. In both of these, the descriptions were concerned with such phrases as "freedom of movement," "blackness of line," "dark forms," "kinesthetic qualities," etc. Nitschke in fact, as reported before, went so far as to assume that he could assess "spontaneity" by subjecting drawings to a blackness count by an IBM machine.[17]

We see that our own factor analysis reveals a distinct relationship between "Handling: Freeness," "Line: Thick," and "Gradation." Some degree of association (.262) also appeared between these criteria and "Texture." Clearly this is a factor that relates tonal-kinesthetic characteristics, and thus sounds very much like that identified by Beittel and Burkhart.

But when we consider the further question of whether these characteristics add up to "Creativity" as such, we can only state that this still seems debateable.

Every study of Creativity places much emphasis on the presence of "originality" or "uniqueness" as an important aspect. Guilford, for example, lists it as one of the four major traits isolated by his battery of tests.[10] We, too, included a criteria called "Originality" in our measure, and yet our factor analysis reveals no apparent relationship, positive or negative, between our "Kinesthetic-Spontaneity" factor and "Originality" (-.023). On the other hand, "Originality" is heavily linked to "Mode: Abstraction," and somewhat less heavily related to "Tension" and "Rhythm." But "Handling: Freeness," the major criteria of the Factor "Kinesthetic-Spontaneity" has a loading of only .038 on the "Conformity" Factor. Consequently the two seem not at all related.

It is important to remember, however, that Beittel and Burkhart's studies were usually based on art products from college art education majors. Our own study utilized a sample of products from several levels: the lower elementary grades through college. It may be that our wider sampling should be expected to produce fewer evidences of creative work than theirs and thus would affect the factor analyses.

In their most recent study, Beittel and Burkhart attempted to learn what other criteria might be present in their construct of "Spontaneity," and by factor analysis showed that these applied: "Detail: Loose," "Media Overlap," "Border of Shapes," "Texture," "Movement Within Shapes," and "Objects: Moving."[2] These cannot really be compared with ours, however, since the list of phenomena with which their analysis started differed considerably from ours.

(6) The relationship of "Balance," "Dominance," and "Unity" shown by the factor analysis confirms the traditional grouping of these terms into the so-called "Principles of Design."

It is interesting to note that the factor we term "Structural/Intellectual" is composed of those phenomena or criteria which can be most consciously manipulated by the artist. Three of them: "Balance," "Dominance," and "Unity" are commonly employed along with "Rhythm" in discussions as the "Principles of Design." Our analysis tends to confirm the validity of this usage with the possible exception of "Rhythm."

It may well be possible that "Rhythm" possesses some subtle qualities quite different from these other three. More research will be needed to confirm or deny this.

(7) A total score was found to discriminate between the "best" and "poorest" products, but did not do so to a significant level.

It was originally proposed that the measure would be scored in such a manner that one general score would result from the entire test. It was hoped that this score could be considered a measure of a previously defined concept, "Primitivism/Sophistication."

In actuality, however, this did not prove to be workable. Although a total score did discriminate between the "best" and "poorest" products as judged by two art education faculty members, the difference was not statistically significant. On the other hand, the scores achieved by these groups of products on one

of our six sub-tests ("Internal Static/Dynamic Action") did show significant differences. This fact, coupled with the results of the factor analysis, seems ample justification for recommending that our measure be scored in terms of each of the sub-tests rather than one general score in future research efforts.

We therefore feel that our proposed title "Primitivism/ Sophistication" is no longer descriptive of the measure and should be discarded.

(8) One sub-test, "Internal Static/Dynamic Action" was shown to significantly discriminate between those products rated "best" and those products rated "poorest."

This sub-test, composed of three criteria: "Symmetry/ Asymmetry," "Non-Rhythm/Rhythm," and "Non-Transparency/ Transparency" was shown to be the only sub-test which significantly ($p = .05$) discriminated between the "best" and "poorest" groups of art products. None of the other tests approached significance, although the mean scores of the "best" products were in most cases higher than the mean scores of the "poorest" products. The "Non-Differentiation/Differtiation" sub-test showed a definite trend toward significance, however.

7. Implications

A. This study can possibly furnish the basis for better communication in this and other related fields. Terminology has been delimited in a reasonably efficient manner which is a definite step upward from the chaos which previously prevailed. While we do not feel that these definitions are necessarily final, we believe that some foundations have been laid for future work in this direction. It should now be possible for workers in this and related fields to be able to describe art products in a much more reliable and valid manner.

B. This study may well serve as a model for others in terms of utilizing statistical techniques to identify and relate criteria and phenomena rather than relying on theoretical orientation or expert opinion. It is believed that the 20 criteria now isolated may be only part of those which might be discovered through similar methods of investigation. Our study, certainly, has displayed the lack of agreement and communication between "authorities" as to what can be discovered or utilized efficiently and points, therefore, to the need for more such inductively-oriented studies in the future.

C. It appears that this study has provided a usable tool for the assessment of specific changes in artistic behavior. This should prove helpful to the researcher who is experimenting with differential treatments of groups as well as to the classroom teacher interested in learning the effects of his or her teaching procedures.

Rather than merely indicating that changes as such have occurred, this measure should make it possible to determine the specific characteristics of the changes made.

D. The finding that one of the six sub-tests ("Internal Static/ Dynamic Action") showed significant differences between those products rated "best" and those rated "poorest" might give impetus to other studies along this line in the future. The investigator believes that this is one of the few occasions where evaluative differences have been linked by statistical techniques to specific kinds of behavior.

Such procedures might well prove to be of value to aestheticians in validating their future statements, which up to this time have rested largely upon personal opinion--a not very trustworthy authority.

E. Our finding that "originality" is not necessarily related to those criteria which form our factor of "Kinesthetic-Spontaneity," but is related to our "abstraction" criteria may well have implications for future research in creativity. As discussed before, others have theorized that spontaneity has some special relationship to creativity, and our results seem to contradict this. This finding appears to demand more research along this line.

In conclusion, may we express our gratitude to the United States Office of Education for the support which has made this study possible, and state hopefully that this support has resulted in some findings of significance to our field. The value of such findings, of course, must always be determined by future usage. Two studies now underway, one in California and one in Illinois, may serve to shed further light on this measure. The fact that this measure finds such immediate acceptance, however, confirms the need for the development of such instruments. We await, with much interest, the results of these two other studies.

Bibliography

1. Beittel, K.R. and Burkhart, R.C. "Strategies of Spontaneous, Divergent, and Academic Art Students," Studies in Art Education, Vol. 5, No. 1, Fall, 1963.

2. Bernheim, G.O. "The Dimensionality of Differential Criteria in the Art Product." Studies in Art Education, Vol. 6, No. 1, Fall, 1964.

3. Bird, M.H. A Study In Aesthetics Cambridge: Harvard University Press, 1932.

4. Burkhart, R.C. Spontaneous and Deliberate Ways of Learning Scranton: International Textbook Co., 1962.

5. Eisner, E.W. "A Typology of Creativity In the Visual Arts," Studies in Art Education, Vol. 4, No. 2, Fall, 1962.

6. Elkisch, Paula "Children's Drawings in a Projective Technique," Psychological Monographs, 1945, 58, No. 1.

7. Goodenough, F.L. Measurement of Intelligence by Drawings New York: Harcourt, Brace, and World, 1926.

8. Goode, W.J. and Hatt, P.K. Methods in Social Research New York: McGraw Hill, 1952.

9. Guilford, J.P. Psychometric Methods New York: McGraw Hill, 1954

10. Guilford, J.P. "Traits of Creativity," in Creativity and Its Cultivation, ed. H.H. Anderson New York: Harper and Brothers, 1959.

11. Harris, D.B. Children's Drawings as Measures of Intellectual Maturity New York: Harcourt, Brace, and World, 1963.

12. Kinget, G.M. The Drawing Completion Test New York: Grune and Shatton, 1952.

13. Kline, L.W. and Carey, G.L. A Measuring Scale for Free-Hand Drawing Baltimore: Johns-Hopkins Press, 1933.

14. Machover, Karen Personality Projection in the Drawings of the Human Figure Springfield: C.C. Thomas, 1949.

15. McFee, J.K. "Creative Problem Solving Abilities of Academically Superior Adolescents" unpublished research report, Secondary Education Project, Stanford University, 1965.

16. Munro, Thomas Scientific Method In Aesthetics New York: W.W, Norton, 1928.

17. Nitschke, Earl "Some Advantages of Using Multiple Drawing Behavior Samples in Lieu of Regular Classwork" Unpublished doctoral dissertation, Pennsylvania State University, 1963.

18. Werner, Heinz A Comparative Psychology of Mental Development Chicago: Follett, 1948.

19. Witkin, H.A., et al Psychological Differentiation New York: John Wiley and Sons, 1962.

20. Wolff, W. "Projective Methods for Personality Analysis of Expressive Behavior in Preschool Children, Character and Personality, 1942, 10.

Table 1
38 Items Submitted to Experts For Rating
In Terms of Possible Efficiency

1. Abstractness
2. Accentuation
3. Alteration of Materials
4. Balance
5. Complexity
6. Distortion
7. Dominance
8. Elaboration
9. Expansion
10. Expressiveness
11. Feeling of Self
12. Fluency of Ideas
13. Gradation of Tone
14. Integration of Forms
15. Integration of Lines
16. Integration of Shapes
17. Light
18. Opposition
19. Pattern
20. Plasticity
21. Proportion
22. Radiation
23. Rhythm
24. Simplicity
25. Spatial Organization
26. Spontaneity
27. Symbolism
28. Technical Skill
29. Tension
30. Texture
31. Transition
32. Transparency
33. Unity
34. Uniqueness
35. Variation in Form
36. Variation in Shape
37. Variation in Use of Line
38. Variation in Use of Materials

INDIANA UNIVERSITY
School of Education
EDUCATION BUILDING
BLOOMINGTON, INDIANA 47405

AREA CODE 812
TEL. NO.

BOB #51-6436
Approval Expires
12-30-65

This letter is written to you to request your expert assistance in a study in which I am attempting to construct a scale suitable for describing several kinds of art products. You are one of twenty-five individuals from the fields of art history, aesthetics or criticism, and art education who have been selected as capable of, and hopefully interested in, rendering this special kind of service. I will be most grateful if you can spare the time to join in this venture.

Your participation in this project is, of course, an extremely important part of the whole study. Expert opinion is recognized as one of the better validation procedures in pursuing studies of this nature.

The project, more fully described in the enclosed paper, is an attempt to establish those items which seem most efficient in the description of various kinds of art products produced by individuals of several age groups. If such items can be isolated and shown to be reliable (which is the next step in the procedure), they will be used to assess the consistency of individual expressive behavior in another study yet to come. Potentially, of course, a descriptive scale such as this might prove to be extremely useful in other problems as well and, if successful, it will be made available for such purposes.

The items you will rate have been isolated from 115 different sources which include previous research studies, the writings of the major aestheticians, critics, and art historians of the past two centuries, and books and articles published on the topic of pure and applied design. This preliminary review of relevant sources revealed a total of 299 possible items. Of these, incidentally, only approximately 16% were repetitive; that is, were used by more than five authorities. This confusion of terms has been recognized as a stumbling block to further progress in all of our fields and many authorities have suggested that empirical studies should be instituted to help to clarify the situation. It is hoped that this study can make a small contribution to this enterprise.

From the larger list of 299 possible items, cited above, some 45 have been selected for your inspection. Each item chosen must have met one of two important criteria: (1) it must have been shown to have a reasonable degree of reliability and validity in previous research studies, or (2) it must have been agreed upon as important by several writers in the books and articles mentioned previously.

Your part in this operation will be to view each item separately and then to place it in a category in terms of its possible utility. The "forced choice" method of rating has been chosen as most efficient for this purpose, and you will be given step-by-step instructions on this procedure (see enclosed "Instruction Sheet"). Please follow the directions explicitly.

Needless to say, your assistance in this matter will be most appreciated. Only through such professional cooperation can we expect to make further progress. As evidence of my personal appreciation, I have provided a place on the "Return Blank" (see enclosures) for you to indicate whether or not you wish to have a copy of the results of this investigation forwarded to you when it is completed. I shall be most happy to do so.

Again, my thanks for your assistance. May I expect a return from you by at the latest?

Sincerely,

Mary J. Rouse
Assistant Professor of Education

MJR:ck

Enclosures:

Copy of study proposal (please keep)
Instruction sheet for rating
38 Item cards
Return blank for your name and address
Stamped, addressed envelope for return

BOB #51-6436
Approval Expires
12-30-65

INSTRUCTION FOR RATING DESCRIPTIVE ITEMS

1. You have exactly 38 descriptive items for which to choose. First read through each of these to grasp the import of the whole.

2. Shuffle all cards together.

3. Next, divide (roughly) all cards into three separate piles in this manner:

   Pile No. 1: In this pile place those cards which you feel are <u>definitely</u> usable in a study of this nature

   Pile No. 2: In this pile place those cards which seem questionable in some respects, but probably not completely unusable

   Pile No. 3: In this pile place those cards which seem <u>definitely not</u> usable

4. Now, subdivide your choices further, in order that your final distribution will conform with the following pattern:

Most Usable ← → Least Usable

Score:	10	9	8	7	6	5	4	3	2	1	0
No. of Cards:	1 card	2 cards	3 cards	4 cards	6 cards	6 cards	6 cards	4 cards	3 cards	2 cards	1 card

[most of these should probably come from your "Definitely usable" (No. 1) pile]

[most of these should probably come from your "Questionable" (No. 2) pile]

[most of these should come from your "Definitely not usable" (No. 3) pile]

5. Following this, pick up each of the cards and in the blank provided, place the "score" number which you see indicated on the table above as belonging to the pile in which the card was placed.

   Thus: the item card that you placed in the pile that rates as absolutely "most usable" should be given the score of "10"

   Another example: The 6 cards which you grouped as belonging in the middle of your "sort" (and thus probably "questionable" in your thinking) should <u>each</u> be given the score of "5"

6. Finally, place all of these cards (have you scored each?) in the envelope provided, together with the blank provided for your name and address and mail as soon as possible. Check on this slip whether or not you wish to receive a copy of the final report of this study when it is completed.

7. Thank you for your invaluable assistance.

Table 4

Results of Rating by Experts in Terms of
Median Scores and Q-Scores
(A Measure of Ambiguity of Ratings)

	Item	Median	Q-Score
1.	Unity	7.7	4.9
2.	Technical Skill	7.0	3.3
3.	Uniqueness	6.8	5.1
4.	Spatial Organization	6.7	4.0
5.	Texture	6.5	3.0
6.	Balance	6.5	3.2
7.	Dominance	6.4	3.3
8.	Variation in Shape	6.3	1.7
9.	Variation in Form	6.3	2.2
10.	Rhythm	6.3	2.8
11.	Variation in Use of Line	6.3	3.0
12.	Simplicity	6.2	2.6
13.	Complexity	6.2	3.2
14.	Accentuation	5.9	2.0
15.	Proportion	5.8	2.1
16.	Integration of Form	5.6	3.3
17.	Elaboration	5.5	1.9
18.	Variety in Use of Materials	5.5	2.9
19.	Transition	5.5	3.4
20.	Integration of Shape	5.4	3.1
21.	Tension	5.3	2.2
22.	Opposition	5.3	2.2
23.	Expressiveness	5.3	6.7
24.	Symbolism	5.0	3.0
25.	Spontaneity	5.0	3.5
26.	Fluency of Ideas	5.0	6.1
27.	Pattern	4.9	1.8
28.	Gradation of Tone	4.9	3.9
29.	Integration of Line	4.7	2.6
30.	Expansion	4.7	2.8
31.	Plasticity	4.7	3.0
32.	Alteration of Materials	4.7	3.5
33.	Transparency	4.5	3.9
34.	Abstractness	4.3	2.7
35.	Distortion	4.3	3.9
36.	Light	4.1	2.6
37.	Feeling of Self	4.0	4.7
38.	Radiation	3.4	4.9

Table 5
Interjudge Correlations, Communalities, Means and Standard Deviations

Criteria	Interrater Correlation (N=3)	Communalities	Mean	S. D.
Non-Differentiation/Differentiation				
Shape: Non-Variation/Variation	.70	.61	2.34	1.09
Mass: Non-Variation/Variation	.68	.77	2.41	0.67
Line: Non-Variation/Variation	.72	.73	2.05	0.80
Amount of Texture	.91	.75	3.01	0.91
Amount of Detail	.86	.80	3.28	0.99
Space				
Flatness/Depth	.84	.79	3.02	0.99
Non-Tension/Tension	.61	.74	3.17	0.77
Non-Symbolism/Symbolism	.71	.74	1.67	0.65
Structural/Intellectual				
Amount of Balance	.63	.86	3.49	0.69
Amount of Unity	.84	.86	3.38	0.92
Non-Dominance/Dominance	.66	.63	3.18	0.94
Non-Craftsmanship/Craftsmanship	.89	.89	3.07	1.07
Kinesthetic/Spontaneity				
Line: Thinness/Thickness	.85	.74	3.39	1.11
Amount of Gradation	.85	.82	2.75	0.99
Handling: Carefulness/Freeness	.85	.88	2.72	1.04
Internally Static/Dynamic Action				
Symmetry/Asymmetry	.63	.65	2.99	0.88
Non-Rhythm/Rhythm	.69	.63	3.57	0.62
Non-Transparency/Transparency	.88	.75	2.83	1.11
Conformity/Non-Conformity				
Non-Originality/Originality	.82	.86	2.52	0.97
Mode: Realism/Abstraction	.83	.87	2.22	1.14

Table 6 Intercorrelations of Criteria

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1. Shapes: Non-V/Var.	1.000	0.568	0.454	0.353	0.358	-0.075	-0.017	-0.038	-0.201	-0.184	-0.245	-0.322	0.430	0.369	0.428	0.283	-0.205	0.196	0.092	-0.355
2. Mass: Non-V/Var.		1.000	0.601	0.625	0.533	0.057	-0.074	-0.129	-0.225	-0.029	0.157	-0.219	0.274	0.308	0.142	0.418	-0.190	0.391	0.179	-0.283
3. Line: Non-V/Var.			1.000	0.567	0.649	-0.068	-0.064	-0.008	-0.322	-0.227	0.010	-0.419	0.315	0.358	0.213	0.268	-0.175	0.314	0.212	-0.172
4. Amt. of Texture				1.000	0.744	-0.086	-0.091	-0.318	-0.031	0.011	0.205	-0.249	0.319	0.544	0.178	0.216	-0.060	0.404	0.041	-0.299
5. Amt. of Detail					1.000	0.208	0.243	0.100	-0.172	-0.160	0.266	-0.264	0.216	0.335	0.230	0.399	0.202	0.649	0.345	0.003
6. Flatness/ Depth						1.000	0.596	0.352	0.300	0.365	0.494	0.344	-0.250	-0.142	-0.151	0.069	0.410	0.442	0.343	0.350
7. Non-Tension/ Tension							1.000	0.384	0.156	-0.028	0.206	0.170	-0.184	-0.113	0.000	-0.022	0.494	0.399	0.514	0.462
8. Non-Sym/ Symbolism								1.000	-0.167	-0.350	-0.013	0.159	-0.172	-0.401	-0.082	-0.059	0.194	0.005	0.292	0.122
9. Amt. of Balance									1.000	0.671	0.362	0.432	0.018	0.019	-0.222	-0.418	0.214	-0.071	0.073	0.121
10. Amt. of Unity										1.000	0.683	0.423	-0.192	0.016	-0.548	-0.112	0.170	0.022	0.017	0.178
11. Non-Crafts/ Craftsmanship											1.000	0.374	-0.255	-0.033	-0.611	0.131	0.487	0.274	0.262	0.357
12. Non-Domin/ Dominance												1.000	-0.294	-0.368	-0.412	-0.187	0.080	-0.019	0.017	0.159
13. Line: Thin/Thickness													1.000	0.421	0.558	0.064	-0.223	0.237	0.176	-0.142
14. Non-Grad/ Gradation														1.000	0.475	0.237	-0.053	0.307	-0.270	-0.388
15. Handling: Car./Freeness															1.000	0.186	-0.179	0.340	-0.017	-0.130
16. Symmetry/ Asymmetry																1.000	0.089	0.544	0.103	0.014
17. Non-Rhy./ Rhythm																	1.000	0.170	0.258	0.501
18. Non-Trans./ Transparency																		1.000	0.362	0.211
19. Non-Orig./Origin.																			1.000	0.576
20. Mode: Realism/Abst.																				1.000

Table 7
Factor Analysis
20 Criteria

	Factor A Non-Differ./ Differentiation	Factor B Space	Factor C Struct./Intell.	Factor D Kines./Spon.	Factor E Internally Static Dynamic Action	Factor F Conformity/ Non-Conformity
Shape: Non-V/Var.	-.581	.142	-.184	.401	.145	.195
Mass: Non-V/Var.	-.868	.002	-.011	.040	-.051	.126
Line: Non-V/Var.	-.804	-.109	-.220	.120	-.006	-.082
Amt. of Texture	-.749	-.212	.173	.262	-.192	.081
Amt. of Detail	-.747	.160	-.054	.174	-.371	-.199
Flatness/Depth	-.059	.722	.388	-.120	-.273	-.172
Non-T/Tension	.054	.702	.079	.071	-.220	-.431
Non-S/Symbolism	.023	.713	-.341	-.282	.136	-.114
Amt. of Balance	.220	.115	.843	.182	.216	-.095
Amt. of Unity	.019	-.110	.897	-.204	-.060	-.023
Non-D/Dominance	.232	.403	.521	-.331	.124	.117
Non-C/Crafts.	-.245	.051	.674	-.414	-.370	-.256
Line: Thinness/Thickness	-.312	-.162	-.054	.715	.276	-.164
Amt. of Gradation	-.290	-.222	.138	.671	-.343	.310

Table 7 (cont.)

	Factor A Non-Differ./ Differentiation	Factor B Space	Factor C Struct./Intell.	Factor D Kines./Spon.	Factor E Internally Static Dynamic Action	Factor F Conformity/ Non-Conformity
Handling: Care./Freeness	-.052	.091	-.424	.827	-.060	.038
Symmetry/ Asymmetry	-.376	-.010	-.260	-.009	-.660	.049
Non-R/Rhythm	.198	.215	.223	-.053	-.541	-.447
Non-T/Trans.	-.436	.348	.071	.304	-.541	-.210
Non-O/Origin.	-.317	.310	.014	-.023	.116	-.807
Mode: Realism/Abstr.	.251	.115	.093	-.161	-.206	-.843

Table 8

Results of T-tests of Mean Scores
Achieved by "Best" and "Poorest" Products

Test	"Best" Products" Mean Scores	"Poorest" Products' Mean Scores	T-test	P
I	14.6	11.6	1.66	n.s.
II	8.0	8.0	.00	n.s.
III	13.2	12.9	.21	n.s.
IV	9.0	8.0	.67	n.s.
V	10.8	8.7	2.21	.05
VI	5.4	4.4	.86	n.s.

Table 9
Final Version of the Measure

Product No. _____
Judge _____

Test I: NON-DIFFERENTIATION/DIFFERENTIATION DOMAIN

1. Shape: Non-Variation/Variation

In regard to shapes (two dimensional areas defined by contour), would you define the product as:

- Shows no variation in shapes; shapes are all similar
- Shapes are mainly similar, but some small variation is shown (approximately 75% to 25%)
- Shows about 50% varied shapes, 50% similar shapes
- Shapes are mainly varied, although some similarity is shown (approximately 75% to 25%)
- Shapes are extremely varied; shows no similarity in shapes

2. Mass: Non-Variation/Variation

In regard to masses (forms having a real or apparent three-dimensional quality), would you describe the product as:

- Shows no variation in masses; masses are all similar
- Masses are mainly similar, but some small variation is shown (approximately 75% to 25%)
- Shows about 50% varied masses, 50% similar masses
- Masses are mainly varied, although some similarity is shown (approximately 75% to 25%)
- Masses are extremely varied; shows no similarity in masses

3. Line: Non-Variation/Variation

In regard to line (the representation of objects, action or decoration by strokes or marks created by any medium, which might include string, wire, thread, as well as brush, pencil, crayon, etc.), would you describe the product as: (Note: Do not include obvious textures as line.)

- Shows no variation in line; lines are all similar
- Lines are mainly similar, but some small variation is shown (approximately 75% to 25%)
- Shows about 50% varied lines, 50% similar lines
- Lines are mainly varied, although some similarity is shown (approximately 75% to 25%)
- Lines are extremely varied; shows no similarity in lines

4. Amount of Texture

In regard to texture (the real or apparent surface quality of an object), would you describe the product as:

- Completely lacking in texture
- Shows small amount of texture
- Shows a moderate amount of texture
- Shows a good deal of texture (at least four different ones)
- Shows a great amount of texture; most surfaces are varied in some way

5. Amount of Detail

In regard to detail (the small, minute parts of the composition), would you describe the product as:

- Shows few or no details; extremely simple
- Shows only small amount of detail
- Shows a moderate amount of detail
- Shows a good deal of detail
- Shows an extremely large amount of detail; highly complex

Test II: SPACE DOMAIN

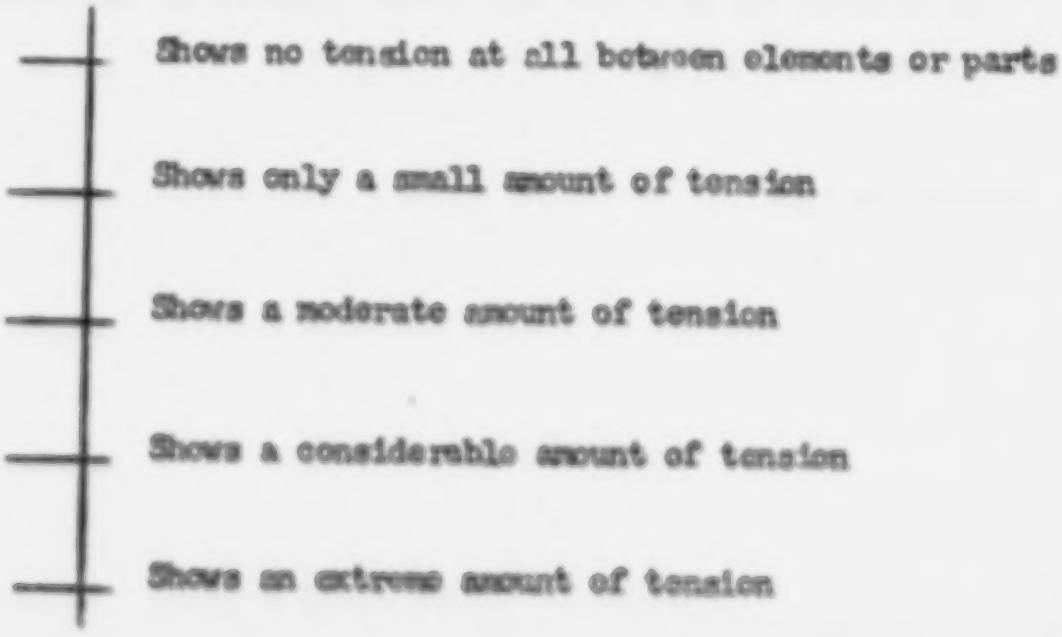
1. Flatness/Depth

In regard to depth (the actual or implied impression of distance inward or backward from the foreground of the composition), would you describe the product as:

- Absolutely flat
- Mostly flat, but shows a small amount of depth
- Shows a moderate amount of depth
- Shows a considerable amount of depth
- Shows many degrees of depth

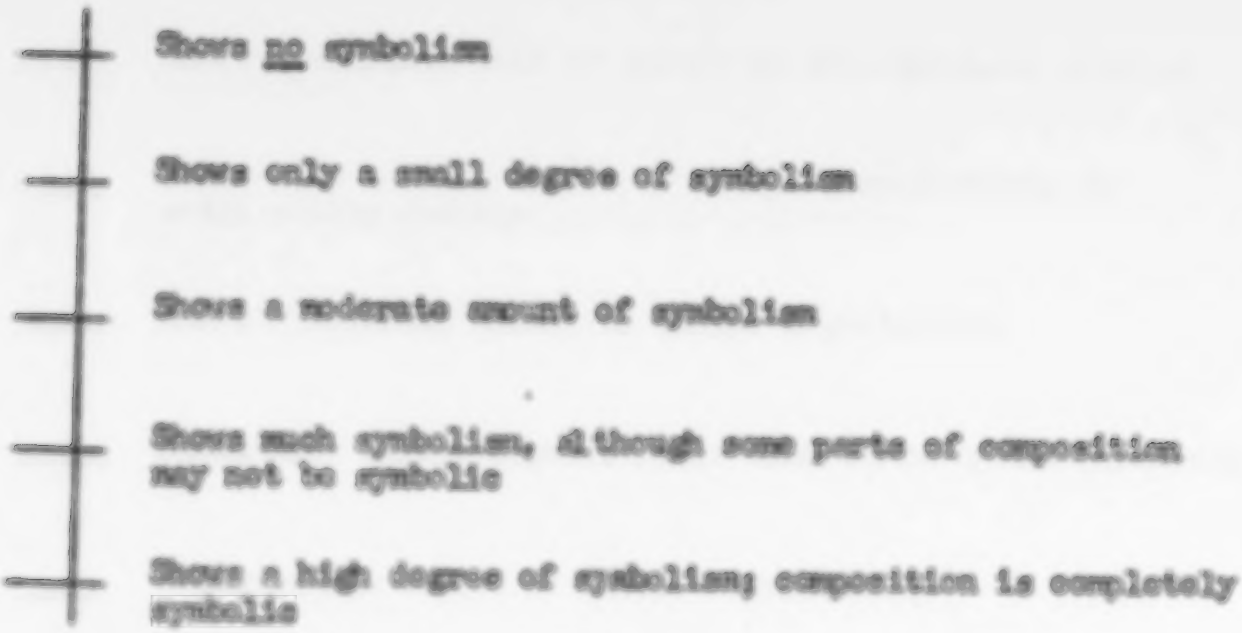
2. Non-Tension/Tension

In regard to tension (actual or implied lines of force set up between the elements or parts of the composition), would you describe the product as:
(Note: Color can also set up tensions)


Shows no tension at all between elements or parts
Shows only a small amount of tension
Shows a moderate amount of tension
Shows a considerable amount of tension
Shows an extreme amount of tension

3. Non-Symbolism/Symbolism

In regard to symbolism (forms which use suggestion to express ideas otherwise intangible), would you describe the product as:


Shows no symbolism
Shows only a small degree of symbolism
Shows a moderate amount of symbolism
Shows much symbolism, although some parts of composition may not be symbolic
Shows a high degree of symbolism; composition is completely symbolic

Test III: STRUCTURAL/INTELLECTUAL DONAIT

1. Amount of Balance

In regard to balance (shown by an equilibrium or poise existing between the elements or parts of the composition, and which may be either symmetrical or asymmetrical in nature), would you describe the product as:

- Shows a complete lack of any type of balance
- Shows only a small amount of any type of balance
- Shows a moderate amount of balance
- Shows a good amount of balance
- Shows a great amount of balance

2. Non-Unity/Unity

In regard to unity (shown by the object's possessing a sense of oneness; i.e., all means and elements appear to be adapted to a single purpose or end), would you describe the product as:

- Shows a complete lack of unity; is disorganized, chaotic
- Shows only a small amount of unity, organization; is still mainly chaotic
- Shows a moderate amount of unity, organization
- Shows more unity, organization, than disunity, disorganization
- Shows a high degree of unity, organization

3. Non-Dominance/Dominance

In regard to dominance (shown by the emphasis of a major idea, theme, or element within the composition; all others play secondary roles), would you describe the product as:

- Shows no dominance at all; all parts are of approximately equal importance
- Shows only a small amount of dominance
- Shows a moderate amount of dominance
- Shows a good deal of dominance
- Shows a great amount of dominance; one idea, theme, or element of much greater importance than the others

4. Non-Craftsmanship/Craftsmanship

In regard to craftsmanship (shown by evidence of competence and skill in the handling of materials and tools which produced the composition), would you describe the product as:

- Shows no craftsmanship; has rough, unfinished, crude appearance
- Shows only a small amount of craftsmanship; is largely rough, unfinished and crude in appearance
- Shows a moderate amount of craftsmanship
- Shows a considerable amount of craftsmanship, finish
- Shows a great amount of craftsmanship; is highly finished in appearance

Test IV: KINESTHETIC-SPONTANEITY DOMAIN

1. Line: Thinness/Thickness

In regard to the weight of line (shown by differences in thickness of lines, ranging from extreme thinness to extreme thickness, would you describe the product as:

Shows extremely thin lines, no thick lines are present

Shows mainly thin lines, although small amount of thicker lines are present (approximately 75% thin, 25% thicker)

Shows approximately 50% thin, 50% thick, or all lines are of moderate thickness

Shows mainly thick lines, although small amount of thin lines are present (approximately 75% thick, 25% thin)

Shows extremely thick lines; no thin lines are present

2. Non-Gradation/Gradation

In regard to gradation (shown by a succession of steps from black through gray to white), would you describe the product as:
(Note: colors can also show gradation)

Shows no gradation

Shows only a small amount of gradation

Shows a moderate amount of gradation

Shows a good deal of gradation

Shows a great deal of gradation

3. Handling: Carefulness/Freeness

In regard to handling (shown by the treatment of materials, and evidence of the kinesthetic character of the utilisation of tools), would you describe the product as:

- Shows an extremely careful, precise treatment of materials or employment of tools
- Shows more of a careful, precise treatment of materials or employment of tools than freeness or fluidness, although some traces of the latter are present
- Approximately 50% careful, precise treatment of materials or employment of tools, 50% freeness or fluidness in treatment of materials or employment of tools
- Shows more freeness, fluidness in treatment of materials or employment of tools than careful, precise treatment or employment, although some traces of the latter are present
- Shows an extremely high degree of freeness, fluidness in treatment of materials or employment of tools

Test V: INTERNAL STATIC/DYNAMIC ACTION DOMAIN

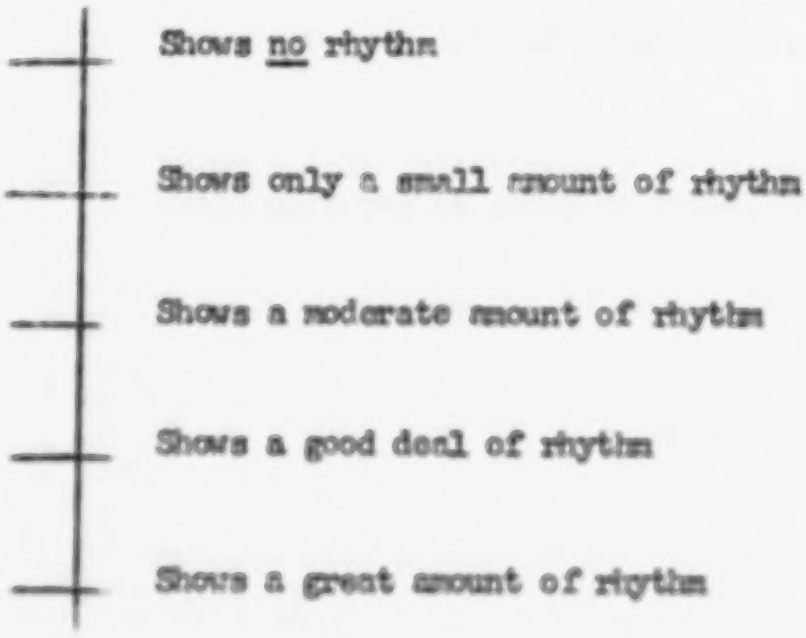
1. Symmetry/Asymmetry

In regard to symmetry/asymmetry (symmetry would be shown by an equal distribution of elements on either side of the central axis; asymmetry would be shown by a control of the opposing elements through a "felt" equilibrium rather than an equal distribution), would you describe the product as:

- Shows complete symmetry
- Shows more symmetry than asymmetry
- About 50% symmetrical, 50% asymmetrical
- Shows more asymmetry than symmetry
- Shows complete symmetry

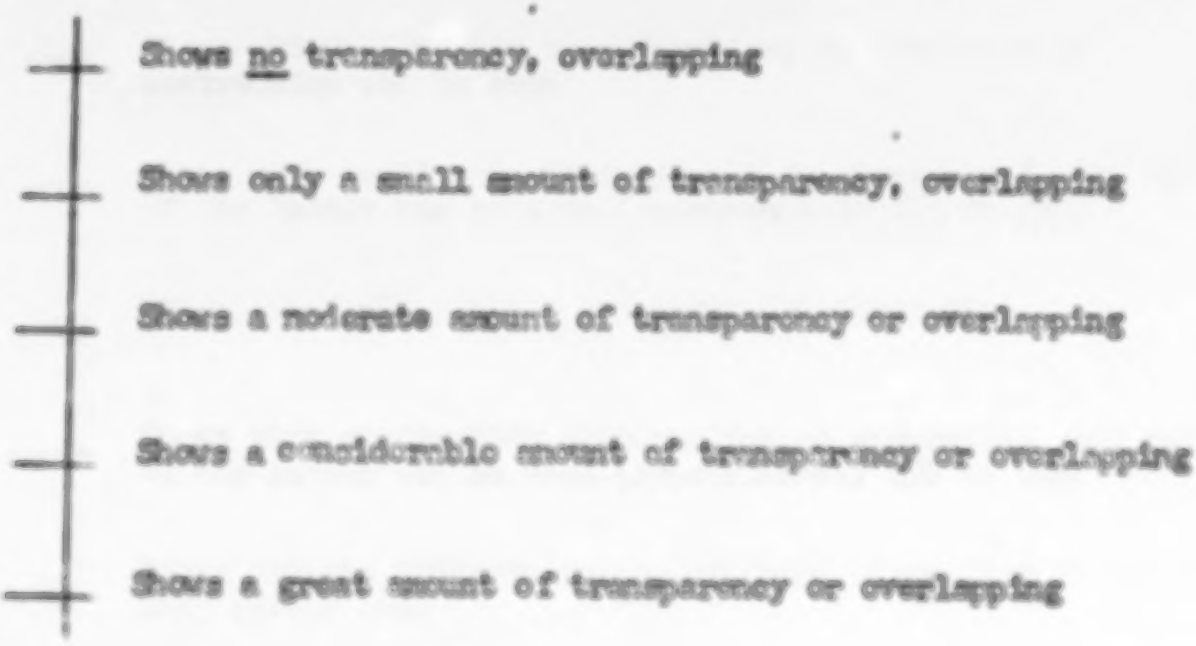
2. Non-Rhythm/Rhythm

In regard to rhythm (shown by a regulated movement, regular repetition or reoccurrence of any element), would you describe the product as:


Shows no rhythm
Shows only a small amount of rhythm
Shows a moderate amount of rhythm
Shows a good deal of rhythm
Shows a great amount of rhythm

3. Non-Transparency/Transparency

In regard to transparency (shown by the actual or apparent interpenetration or overlapping of planes,. lines and/or colors, one or more of which can be observed behind or under another), would you describe the product as:
(Note: this item can also apply to constructions or sculpture where string, wire, etc., might be seen to overlap or penetrate)


Shows no transparency, overlapping
Shows only a small amount of transparency, overlapping
Shows a moderate amount of transparency or overlapping
Shows a considerable amount of transparency or overlapping
Shows a great amount of transparency or overlapping

Test VI: CONFORMITY/NON-CONFORMITY DOMAIN

1. Non-Originality/Originality

In regard to originality (shown by inventiveness, uniqueness, non-imitativeness), would you describe the product as:

- Shows complete lack of originality; is imitative, stereotyped
- Shows mainly imitativeness or stereotypy; although some small part may be somewhat out-of-the-ordinary
- Shows moderate amount of originality
- Shows a good deal of originality
- Shows a high degree of originality

2. Mode of Representation: Realism/Abstraction

In regard to the mode of representation employed, either realism or abstraction (realism would be shown by the depiction or attempt to achieve depiction of objects as they exist; abstraction would be shown by the attempt to generalize or universalize the object), would you describe the object as:
(Note: here you are looking for a mode created by intent, and not merely by a lack of skill)

- Shows extreme realism or naturalism; no evidence of abstraction can be seen
- Shows more realism or naturalism than abstraction, but some of the latter can be seen (approximately 75% to 25%)
- Shows about 50% realism and about 50% abstraction
- Shows more abstraction than realism or naturalism, but some of the latter can be seen (approximately 75% to 25%)
- Shows extreme abstraction; no evidence of realism or naturalism can be seen